ESTA TESE FOI JULGADA ADEQUADA PARA A OBTENÇÃO DO TÍTULO DE

"MESTRE EM CIÊNCIAS"

ESPECIALIDADE EM MATEMÁTICA E APROVADA EM SUA FORMA FINAL PELO CURSO DE PÓS-GRADUAÇÃO

Prof. INDER JEET TANEJA, Ph.D.

Coordenador

BANCA EXAMINADORA:

Prof. GUR DIAL, Ph.D.

Orientador

Prof. INDER JEET TANEJA, Ph.D.

Prof. LEON R. SINAY, Ph.D.

MEDIDAS DE INFORMAÇÃO
E A
DESIGUALDADE DE INDEPENDÊNCIA

por Margley M. Moura

Orientador: Prof. GUR DIAL, Ph.D

- Julho/82 -

OFEREÇO

A Ireide, minha mulher
A meus filhos Vicente
Eduardo e Margley Junior
pela dedicação constante.

DEDICO

A Vicente Moura, meu pai (IN MEMORIAN)
A Toinha Moura, minha mãe
pela compreensão e ajuda
em todas as etapas de
minha vida

AGRADECIMENTOS

Sou profundamente grato ao Prof. Gur Dial, Ph.D, pela maneira precisa e paciente com a qual contribuiu com os seus ensinamentos em todas as fases de preparação deste trabalho.

Gostaria de agradecer particularmente ao Prof. Leon R. Sinai, Ph.D., pelas várias críticas e sugestões que resultaram em um vasto número de melhoramentos e modificações. Sou também grato ao Prof. Inder Jeet Taneja, Ph.D que leu cuidadosamente a redação final deste trabalho.

Também agradeço a José Herisberto de Medeiros Miranda pela rapidez e perfeição de seu trabalho datilográfico.

Margley M. Moura

Florianópolis, SC

Julho, 1982.

RESUMO

Nsste trabalho, apresentamos a desigualdade de independência, sua motivação e importância no campo da Teoria da Informação.

Na literatura da Teoria da Informação existem diferentes medidas de informação, tais como: entropia de Shannon, entropia de tipo α, tipo (α, β), ordem α, ordem (α, β) e a entropia de Mittal.

A validade desta desigualdade foi testada com as supra citadas medidas, particularmente com a entropia de Mittal.

ABSTRACT

The independence inequality is presented considering its motivation and importance to the Information Theory.

According to the Information Theory literature, there exist different measures of information such as Shannon's entropy, Mittal's entropy, and entropies of type α, type (α, β), order α, and order (α, β).

The validity of this inequality was tested with those mentioned measures, particularly with Mittal's entropy.

Í N D I C E

Página

CAPÍTULO I

CONCEITOS E RESULTADOS FUNDAMENTAIS

1.1. INTRODUÇÃO

A idéia de informação já nos é bastante familiar e, como sabemos, esta é chegada até nós através dos diversos meios de comunicações de que dispomos, tais como: jornais, periódicos especializados, revistas, rádio, televisão, telefone, e até mesmo da simples conversa que temos no dia-a-dia.

A Teoria da Informação, de uma maneira geral, está intimamente relacionada com os problemas de origem matemática surgidos em conexão com o armazenamento, transformação e transmissão de informações.

As primeiras investigações nesse campo da ciência foram feitas por Nyquist [15] em 1924 e por Nyquist [16] e Hartley [12] em 1928.

Em 1948, Claude Shannon, com a publicação do seu famoso trabalho intitulado "A Mathematical Theory of Comunication" (Bell System Technical Journal), introduziu um modelo matemático para medir o valor médio da informação fornecida por um experimento. Por sua analogia com a quantidade introduzida por Boltzmann na teoria cinética dos gases, foi chamada de entropia (entropia de Shannon), a qual será apresentada no parágrafo 1.3.

Com essa valiosa contribuição de Shannon à Teoria da Informação, esta passou a ser uma atrativa área de pesquisa e, por consequência, passou a ser aplicada nos mais diferentes ramos do conhecimento humano, tais como: Termodinâmica [3] , Economia [26], Biologia [18], Cibernética [28], etc.

A Teoria da Informação sob a forma clássica de Wiener e Shannon baseia-se na noção de probabilidade.

1.2. DEFINIÇÕES E NOTAÇÕES

Nessa seção nosso objetivo primordial consiste em introduzir algumas definições, bem como algumas notações, sem as quais não poderíamos dar prosseguimento aos nossos objetivos no presente trabalho. Aqui, também, são introduzidos alguns resultados de relevância fundamental para os nossos propósitos.

Sejam $X = \{x_1, x_2, \ldots, x_n\}$ e $Y = \{Y_1, Y_2, \ldots, Y_m\}$ duas variáveis aleatórias discretas com distribuições de probabilidades dadas, respectivamente, por $P=(p(x_1), p(x_2), \ldots, p(x_n))$ e $Q = (p(y_1), p(y_2), \ldots, p(y_m))$, tais que:

$$p(x_i) \geq 0, \quad p(y_j) \geq 0, \; i=1,\ldots, n; \; j=1,\ldots, m;$$

$$\sum_{i=1}^{n} p(x_i) = \sum_{j=1}^{m} p(y_j) = 1.$$

Suponhamos que X e Y estão associadas a um mesmo experimento. Então X e Y têm probabilidade conjunta dada por

$$p(x_i,y_j) = p\,\{\, X=x_i,\; Y=y_j \,\},$$

para $i=1,\ldots,n$ e $j=1,\ldots,m$.

O conjunto de definições que damos a seguir é de fundamental importância.

Definição 1.2.1

A probabilidade $p(x_i)$ do evento $X=x_i$ no primeiro experimento indiferente a do segundo, é definida pela relação

$$p(x_i) = \sum_{j=1}^{m} p(x_i,y_j) \qquad (1.2.1)$$

para $i=1,\ldots,n$.

Definição 1.2.2

A probabilidade $p(y_j)$ do evento $Y=y_j$ no segundo experimento, indiferente a do primeiro, definimos pela relação

$$p(y_j) = \sum_{i=1}^{n} p(x_i,y_j) \qquad (1.2.2)$$

para $j=1,\ldots,m$.

Definição 1.2.3

A probabilidade condicional, $p(x_i/y_j)$, de ocorrer o evento x_i do experimento X, dado que o evento y_j do experimento Y ocorreu, é definida por

$$p(x_i/y_j) = \frac{p(x_i,y_j)}{p(y_j)} \qquad (1.2.3)$$

para $i=1,\ldots,n$; $j=1,\ldots,m$.

Definição 1.2.4

A probabilidade condicional, $p(y_j/x_i)$, de ocorrer o evento y_j do experimento Y, dado que o evento x_i do experimento X ocorreu, definimos pela relação

$$p(y_j/x_i) = \frac{p(x_i, y_j)}{p(x_i)} \tag{1.2.4}$$

para $i=1,\ldots,n$; $j=1,\ldots,m$.

Se os experimentos X e Y são independentes, então valem as relações

$$p(x_i/y_j) = p(x_i)$$

e

$$p(y_j/x_i) = p(y_j),$$

já que neste caso $p(x_i, y_j) = p(x_i).p(y_j)$.

Daqui por diante, por simplicidade de notação, convencionaremos em usar as notações abreviadas: $p(x_i) = p_i$, $p(y_j) = q_j$, $p(x_i, y_j) = \P_{ij}$ e, finalmente, $p(y_j/x_i) = q_{ij}$. Assim, podemos escrever as três seguintes relações:

$$p_i = \sum_{j=1}^{m} \P_{ij}, \quad q_j = \sum_{i=1}^{n} \P_{ij}$$

e

$$\P_{ij} = p_i q_{ij}$$

para $i=1,\ldots,n$ e $j=1,\ldots,m$.

Definição 1.2.5

Um canal discreto sem memória com alfabeto de entrada $X=\{x_1, x_2,\ldots, x_n\}$ e alfabeto de saída $Y=\{y_1, y_2,\ldots, y_m\}$, onde X é uma variável aleatória discreta com distribuição de probabilidade $P=(p_1, p_2,\ldots, p_n)$, Y uma variável aleatória discreta com distribuição de probabilidade $Q=(q_1, q_2,\ldots, q_m)$, é definido pela matriz de ordem n x m abaixo:

$$\mathcal{Q} = (q_{ij})_{(n,m)} \quad (i=1,\ldots,n;\ j=1,\ldots m) \text{ com } q_{ij} \geq 0 \text{ e}$$

$$\sum_{j=1}^{m} q_{ij} = 1 \quad (i=1,\ldots,n).$$

A matriz acima é cumumente chamada de matriz de transição ou matriz canal ou, ainda, matriz de probabilidade condicional.

Denotamos por Γn o conjunto de todas as n-uplas de distribuições de probabilidade $P=(p_1, p_2, \ldots, p_n)$, as quais satisfazem as condições seguintes:

$$p_i \geqq 0, \quad i=1,\ldots,n$$

$$\sum_{i=1}^{n} p_i = 1.$$

Tal conjunto é também denominado de "Conjunto de todas as n-uplas de distribuições de probabilidade finitas e completas". Em forma abreviada, escrevemos

$$\Gamma_n = \{(p_1,p_2,\ldots,p_n); \quad \sum_{i=1}^{n} p_i=1, \; p_i \geqq 0, \; i=1,\ldots,n\} \qquad (n=2,3,\ldots).$$

Usaremos também, aqui, o símbolo $\mathbb{R}$ para denotar o conjunto de todos os números reais.

Para finalizar esta seção temos, ainda, a seguinte

Definição 1.2.6

Seja $P=(p_1,p_2,\ldots,p_n)$ uma distribuição de n probabilidades, e seja $Q=(q_1,q_2,\ldots,q_m)$ uma distribuição de m probabilidades. A distribuição produto entre P e Q de n.m probabilidades, aqui denotada por P^*Q, é definida por

$$P^*Q = (p_1q_1,p_1q_2,\ldots,p_1q_m,p_2q_2,\ldots,p_2q_m;\ldots;p_nq_1,\ldots,\ldots,p_nq_m)$$

1.3. ENTROPIA DE SHANNON. DEFINIÇÃO E PROPRIEDADES

Apresentamos, nessa seção, a definição da entropia de Shannon mencionada no parágrafo 1.1. Após um breve comentário sobre tal entropia, daremos uma lista de propriedades satisfeitas por esta medida de informação, que embora não completa é de fundamental importância, pois este conjunto de propriedades nos dá, de modo significativo, o comportamento desta medida.

Seja $X = \{x_1,\ldots,x_n\}$ uma variável aleatória discreta com distribuição de probabilidade $P = (p_1,\ldots,p_n) \in \Gamma_n$. Assim, temos a seguinte

Definição 1.3.1

A sequência de funções $H_n : \Gamma_n \to \mathbb{R}$ $(n = 2,3,\ldots)$, definida pela relação

$$H_n(P) = H_n(p_1,\ldots,p_n) = -\sum_{i=1}^{n} p_i \log(p_i), \tag{1.3.1}$$

é chamada de entropia de Shannon ou simplesmente entropia ou, ainda, incerteza média.

Na expressão (1.3.1) "log" denota o logarítimo e, em virtude da base do sistema de codificação das possíveis mensagens que a variável pode e-mitir, ser arbitrária, a base de "log" no cálculo da entropia também o será. Por exemplo, a base de log em (1.3.1) será "2" se estivermos usando base binária na codificação.

Em vista de

$$\lim_{x\to 0^+} (x.\log(x)) = 0,$$

convencionaremos que $0.\log(0) = 0$.

A unicidade de (1.3.1), no sentido de que somente H_n satisfaz o conjunto de propriedades que a apresentaremos a seguir, foi provada por Shannon utilizando um razoável conjunto de axiomas. Sobre esta medida de informação, podemos dizer ainda, que ela tem sido caracterizada de várias maneiras por diversos autores [1]. Finalmente, vale salientar que H_n é uma função real e diferenciável, exceto nos vértices de Γ_n.

A seguir fazemos a apresentação de algumas propriedades satisfeitas pela entropia de Shannon. As respectivas demonstrações serão omitidas, mas podem ser vistas, por exemplo, na referência [1].

I. PROPRIEDADES ALGÉBRICAS DE H_n:

(i) SIMETRIA:

$$H_n(p_1,\ldots p_n) = H_n(p_{\sigma(1)},\ldots,p_{\sigma(n)}),$$

para toda distribuição $(p_1,\ldots p_n) \in \Gamma_n$, onde $\sigma : \{1,\ldots,n\} \to \{1,\ldots,n\}$ é uma permutação.

Esta propriedade, em palavras, quer dizer simplesmente: a informação obtida de um experimento é invariante quanto a ordem dos eventos.

(ii) NORMALIDADE:

$$H_2(\frac{1}{2}, \frac{1}{2}) = 1.$$

(iii) EXPANSABILIDADE:

$$H_n(p_1,\dots,p_n) = H_{n+1}(0,p_1,\dots,p_n) = \dots$$

$$= H_{n+1}(p_1,\dots,p_i,0,p_{i+1},\dots,p_n)$$

$$= H_{n+1}(p_1,\dots,p_n,0)$$

$$(i = 1,\dots,n-1).$$

Em palavras, a informação obtida de um experimento é a mesma se a este é adicionado um evento com probabilidade nula.

(iv) DECISIVIDADE:

$$H_2(1,0) = H_2(0,1) = 0;$$

Isto é, nenhuma informação é obtida de um experimento com dois resultados, sendo um com probabilidade 1 e o outro com probabilidade 0 (zero).

(v) ADITIVIDADE FORTE:

$$H_{nm}(p_1q_{11},\dots,p_1q_{1m};p_2q_{21},\dots,p_2q_{2m};\dots;p_nq_{n1},\dots,p_nq_{nm})$$

$$= H_n(p_1,\dots,p_n) + \sum_{i=1}^{n} p_1H_m(q_{i1},\dots,q_{im}),$$

para todo $(p_1,\dots,p_n) \varepsilon \Gamma_n$ e $(q_{i1},\dots,q_{im}) \varepsilon \Gamma_m$, $i=1,\dots,n$;

Quer dizer, a informação obtida de dois experimentos é igual a informação obtida do primeiro experimento mais a informação condicional do segundo experimento, com respeito ao primeiro.

(vi) ADITIVIDADE:

$$H_{nm}(p_1q_1,\dots,p_1q_m;p_2q_1,\dots,p_2q_m;\dots;p_nq_1,\dots,p_nq_m)$$

$$= H_n(p_1,\dots,p_n) + H_m(q_1,\dots,q_m)$$

para todo $n,m \geqq 2$, $(p_1,\dots,p_n) \in \Gamma_n$, $(q_1,\dots q_m) \in \Gamma_m$.

Em palavras, esta propriedade diz que a informação obtida da combinação de dois experimentos independentes é igual a soma das suas informações, quando de modo isolados.

(vi) RECURSIVIDADE:

$$H_n(p_1,\dots,p_n) = H_{n-1}(p_1+p_2,p_3,\dots,p_n) + \\ + (p_1+p_2)H_2\left(\frac{p_1}{p_1+p_2}, \frac{p_2}{p_1+p_2}\right)$$

para todo $(p_1,\dots p_n) \in \Gamma_n$, no qual $p_1+p_2 > 0$.

II. PROPRIEDADES ANALÍTICAS DE H_n:

(i) NÃO-NEGATIVIDADE:

$$H_n(p_1,\dots,p_n) \geqq 0$$

para todo $n=1,2,\dots$ e para todo $(p_1,\dots,p_n) \in \Gamma_n$.

(ii) MAXIMILIDADE:

$$H_n(p_1,\dots p_n) \leqq H_n\left(\frac{1}{n},\dots, \frac{1}{n}\right) \\ = \log(n)$$

para toda distribuição $(p_1,\dots,p_n) \in \Gamma_n$, $(n=2,3,\dots)$.

Corolário

$$H_2(p_1,p_2) \leqq 1,$$

para todo $(p_1,p_2) \in \Gamma_n$.

(iii) SUB-ADITIVIDADE:

$$H_{nm}(\P_{11},\P_{12},\ldots,\P_{1m};\P_{21},\P_{22},\ldots;\P_{2m};\ldots;\P_{n1},\P_{n2},\ldots,\ldots,\ldots,\P_{nm})$$

$$\leqq H_n\left(\sum_{j=1}^{m}\P_{1j},\ \sum_{j=1}^{m}\P_{2j},\ldots,\ \sum_{j=1}^{m}\P_{nj}\right) +$$

$$+ H_m\left(\sum_{i=1}^{m}\P_{i1},\ \sum_{i=1}^{m}\P_{i2},\ldots,\ \sum_{i=1}^{m}\P_{im}\right),$$

para todo $n \geqq 2$, $m \geqq 2$, $(\P_{11},\ldots,\P_{1m};\ldots;\P_{n1},\ldots,\P_{nm}) \in \Gamma_{nm}$.

Em palavras, a informação obtida da combinação de dois experimentos (não necessariamente independentes) não é maior do que a soma das informações, quando de modos isolados.

(iv) CONCAVIDADE:

A entropia H_n, definida em (1.3.1), é côncava; quer dizer:

$$\lambda H_n(P) + \mu H_n(Q) \leqq H_n(\lambda P + \mu Q),$$

para todo $\lambda,\mu \in \mathbb{R}$ tais que $\lambda + \mu = 1$ e distribuições $P,Q, \in \Gamma_n$.

(v) CONTINUIDADE:

A função $H_n : \Gamma_n \to \mathbb{R}$ é uma função contínua de P_1, $P_2,\ldots,P_n$.

(vi) Se $(P_1,\ldots,P_n) \in \Gamma_n$ e $(q_1,\ldots,q_m) \in \Gamma_m$, então

$$H_n(P_1,\ldots,P_n) \leqq -\sum^{n} p_i \log(q_i),$$

com a igualdade acontecendo se, e somente se, $p_i = q_i$, $i=1,\ldots,n$.

As propriedades analíticas têm a sua importância evidenciada a partir da necessidade de adaptar um fenômeno em estudo, ao modelo da entropia, enquanto as propriedades algébricas suportam o aspecto computacional do mesmo modelo

1.4. GENERALIZAÇÕES DA ENTROPIA

Reservamos este parágrafo para fazer a apresentação de algumas generalizações da entropia de Shannon. Consideramos diferentes medidas, incluindo as entropias de Rényi (entropia de ordem α), entropia de Daróczy (entropia de grau α), entropia de Sharma-Taneja (Entropia de grau (α, β)), entropia de ordem (α,β), entropia de Mittal (entropia paramétrica generalizada ou entropia de ordem α e tipo β) e, por fim, a medida de informação R-Norm.

Após as definições destas medidas de incertezas, apresentamos uma lista de propriedades satisfeitas por tais entropias. A exemplo do parágrafo anterior, aqui também, omitiremos as respectivas demonstrações.

1.4.1. - ENTROPIA DE ORDEM α

Tal entropia foi desenvovida por Rényi em 1961 e, por isto é denominada também de entropia de Rényi. As propriedades dessa medida de informação, bem como sua caracterização, foram estudadas recentemente por Aczél e Daróczy [1].

Definição (entropia de ordem α)

A entropia de ordem $\alpha \neq 1$ de uma distribuição de probabilidade $P = (p_1,\ldots,p_n) \in \Gamma_n$ $(n=2,3,\ldots)$, é definida por

$$(1.4.1) \qquad {}_{\alpha}H_n(P) = {}_{\alpha}H_n(p_1,\ldots p_n) = \frac{1}{1-\alpha}\log\left(\sum_{i=1}^{n} p_i^{\alpha}\right),$$

onde usamos a definição $0^{\alpha} := 0$, para todo $\alpha \in \mathbb{R}$, $\alpha > 0$.

É fácil ver que quando α tende para 1, a entropia de Rényi tende para a entropia de Shannon; isto é,

$$\lim_{\alpha\to 1} {}_{\alpha}H_n(P) = H_n(P),$$

onde $P \in \Gamma_n$. Por esta razão, diz-se também que a entropia de Shannon é a entropia de ordem 1.

A seguir, damos uma lista de algumas propriedades satisfeitas por ${}_{\alpha}H_n$. Aqui, também, dividimos tais propriedades em dois grupos, a saber: um formado pelas propriedades de caráter algébrico e o outro constituído de propriedades de caráter puramente analítico.

I. PROPRIEDADES ALGÈBRICAS DE ${}_{\alpha}H_n$:

A entropia de Rényi satisfaz as seguintes propriedades:

(i) ${}_{\alpha}H_n(p_1,\ldots,p_n)$ é uma função simétrica de $P_1,\ldots,p_n$;

(II) ${}_{\alpha}H_n(p_1,\ldots,p_n)$ é expansável;

(III) ${}_{\alpha}H_n(p_1,\ldots,p_n)$ é aditiva;

(iv) ${}_{\alpha}H_n$ é decisiva: ${}_{\alpha}H_2(1,0) = {}_{\alpha}H_n(0,1) = 0$;

(v) ${}_{\alpha}H_n(p_1,\ldots,p_n)$ é não-recursiva.

II. PROPRIEDADES ANALÍTICAS DE ${}_{\alpha}H_n$:

A entropia de Rényi satisfaz, também, as seguintes propriedades:

(i) ${}_{\alpha}H_n(p_1,\ldots,p_n) \geqq 0$; i.e, ${}_{\alpha}H_n$ é não-negativa;

(ii) ${}_{\alpha}H_n(1/2,\ 1/2) = 1$,

(iii) ${}_{\alpha}H_n(p_1,\ldots p_n) \leqq {}_{\alpha}H_n(1/n,\ldots,1/n)$
$= \log(n)$,

para toda distribuição $(p_1,\ldots,p_n) \in \Gamma_n$,

(iv) ${}_{\alpha}H_n(p_1,\ldots,p_n)$ é contínua em ;

(v) ${}_{\alpha}H_n(p_1,\ldots,p_n)$ é uam função estritamente decrescente de α;

(vi) ${}_{\alpha}H_n(1/n,\ldots,1/n)$ é uma função monótona crescente de n;

(vii) ${}_{\alpha}H_n(p_1,\ldots,p_n)$ é côncavo para $0 < \alpha \leqq 1$, se $n \geqq 2$.

1.4.2. - ENTROPIA DE GRAU α

Daróczy (1970), introduziu o conceito de uma função de informação de grau α e definiu a entropia de grau α para uma distribuição de probabilidade $P = (p_1,\ldots,p_n)$ de uma variável aleatória $X = \{x_1,\ldots,x_n\}$. Explicitamente tal entropia é dada abaixo, na seguinte

Definição (entropia de grau α)

A entropia de grau $\alpha \neq 1$ de uma distribuição de probabilidade $P = (p_1,\dots,p_n) \in \Gamma_n$ $(n=2,3,\dots)$, é definida por

$$(1.4.2) \quad H_n^{\alpha}(p_1,\dots,p_n) = (2^{1-\alpha}-1)^{-1}\left(\sum_{i=1}^{n} p_i^{\alpha} - 1\right).$$

Aqui também, usamos a definição $0^{\alpha} := 0$, para todo $\alpha \in \mathbb{R}$, $\alpha > 0$.

Tal entropia é também denominada de entropia de Daróczy.

Quando $\alpha \to 1$, temos

$$\lim_{\alpha\to 1} H_n^{\alpha}(p_1,\dots,p_n) = H_n(p_1,\dots,p_n)$$

$$= -\sum_{i=1}^{n} p_i \log(p_i),$$

a qual é a entropia de Shannon.

As propriedades, bem como a caracterização dessa entropia, foram estudadas por Taneja [24], [25], e Aczél e Daróczy [1]. Algumas dessas propriedades são:

(i) $H_n^{\alpha}(p_1,\dots,p_n)$ é uma função simétrica de $p_1,\dots p_n$;

(ii) $H_n^{\alpha}(p_1,\dots,p_n)$ é recursiva; i.e,

$$H_n^{\alpha}(p_1,\dots,p_n) = H_{n-1}^{\alpha}(p_1+p_2,p_3,\dots,p_n) + (p_1+p_2)\cdot H_2^{\alpha}\left(\frac{p_1}{p_1+p_2},\frac{p_2}{p_1+p_2}\right),$$

onde $(p_1,\dots,p_n) \in \Gamma_n$ e $p_1+p_2 > 0$.

(iii) $H_n^{\alpha}(p_1,\dots,p_n)$ é expansível; quer dizer,

$$H_n^{\alpha}(p_1,\dots,p_n) = H_{n+1}^{\alpha}(p_1,p_2,\dots,p_n,0).$$

(iv) $H_n^{\alpha}(p_1,\dots,p_n)$ satisfaz a desigualdade

$$H_n^{\alpha}(p_1,\dots,p_n) \leq H_n^{\alpha}(1/n,\dots,1/n),$$

com a igualdade acontecendo se, e somente se, $p_i=1/n$, para todo $i=1,\ldots,n$ e $(p_1,\ldots,p_n) \in \Gamma_n$.

(v) $H_n^\alpha(p_1,\ldots,p_n)$ é não-aditiva; i.e,

$$H_{nm}^\alpha(P\,Q) = H_n^\alpha(P) + H_n^\alpha(Q) + A_\alpha H_n^\alpha(P)H_m^\alpha(Q),$$

onde $P = (p_1,\ldots,p_n) \in \Gamma_n$, $Q = (q_1,\ldots,q_m) \in \Gamma_m$ e $A_\alpha := (2^{1-\alpha}-1)^{-1}$.

(vi) $H_n^\alpha(1/n,\ldots,1/n)$ é uma função monótona de n.

(vii) $H_n^\alpha(p_1,\ldots,p_n)$ é sub-aditiva.

1.4.3 - ENTROPIA DE GRAU (α,β)

A entropia de grau (α,β), $\alpha \neq 1$, $\beta \neq 1$, é significativamente uma nova medida de incerteza. Aqui, nosso objetivo para com esta medida de informação, a exemplo do que vimos fazendo, consiste em defini-la e, em seguida, listar algumas de suas mais importantes propriedades.

Definição (entropia de grau (α,β))

A entropia de grau $(\alpha\ \beta)$, $\alpha \neq 1$, $\beta \neq 1$, (entropia de Sharma-Taneja), de uma distribuição de probabilidade $P = (p_1,\ldots,p_n) \in \Gamma_n$, denotada por $H_n^{\alpha,\beta}(P)$, definimos da maneira seguinte:

$$H_n^{\alpha,\beta}(p_1,\ldots,p_n) = (2^{1-\alpha} - 2^{1-\beta})^{-1} \cdot \sum_{i=1}^{n} (p_i^\alpha - p_i^\beta), \tag{1.4.3}$$

$\alpha \neq \beta$, $\alpha,\beta > 0$.

Aqui também usamos a definição $0^\gamma := 0$, para todo γ.

Observamos que a entropia de grau (α,β) é uma função real de Γ_n em $\mathbb{R}$, onde $n=2,3,\ldots$ E mais, que $H_n^{\alpha,1}(P) = H_n^\alpha(P)$.

Daremos agora algumas das propriedades algébricas e analíticas da entropia $H_n^{\alpha,\beta}$. Primeiro consideramos as propriedades de caráter puramente algébrico. Estas são:

(i) $H_n^{\alpha,\beta}(p_1,\ldots,p_n)$ é simétrica com respeito a $p_1,\ldots,p_n$;

(ii) $H_n^{\alpha,\beta}(p_1,\ldots,p_n)$ é expansível; isto é,

$$H_n^{\alpha,\beta}(p_1,\ldots,p_n) = H_{n+1}^{\alpha,\beta}(p_1,\ldots,p_n,0);$$

(iii) $H_n^{\alpha,\beta}(p_1,\ldots,p_n)$ é não recursiva;

(iv) $H_n^{\alpha,\beta}$ é decisiva: $H_2^{\alpha,\beta}(1,0) = H_n^{\alpha,\beta}(0,1)$;

Algumas das propriedades analíticas satisfeitas por $H_n^{\alpha,\beta}$ são:

(i) $H_n^{\alpha,\beta}(p_1,\ldots p_n) \geq 0$;

(ii) $H_n^{\alpha,\beta}(p_1,\ldots,p_n)$ é uma função de suas variáveis;

(iii) $H_n^{\alpha,\beta}(1/2,1/2) = 1$.

1.4.4 - ENTROPIA DE ORDEM (α,β)

A entropia de ordem (α,β) foi introduzida por Aczél por volta de 1963. Esta medida de incerteza é definida como segue:

Definição (entropia de ordem (α,β))

A entropia de ordem (α,β) de uma distribuição de probabilidade $P = (p_1,\ldots,p_n) \in \Gamma_n$, aqui denotada por ${}_{\alpha,\beta}H_n$, é definida por

(1.4.4) $$ {}_{\alpha,\beta}H_n(p_1,\ldots,p_n) = \frac{1}{\beta-\alpha} \log \left(\frac{\sum_{i=1}^{n} p_i^{\alpha}}{\sum_{i=1}^{n} p_i^{\beta}} \right), \ \alpha \neq \beta,\ \alpha \neq 1,$$

$\beta \neq 1$, $\alpha,\beta > 0$, com a definição $0^{\gamma} := 0$, para todo $\gamma \in \mathbb{R}$.

Evidentemente que se tomarmos $\beta = 1$ em (1.4.4), esta se reduzirá à entropia de Rényi; isto é, temos que

$$ {}_{\alpha,1}H_n(p_1,\ldots,p_n) = {}_{\alpha}H_n(p_1,\ldots,p_n).$$

Algumas propriedades satisfeitas por ${}_{\alpha,\beta}H_n$ são:

(i) ${}_{\alpha,\beta}H_n(p_1,\ldots,p_n)$ é simétrica com respeito a $p_1,\ldots,p_n$;

(ii) ${}_{\alpha,\beta}H_2(1,0) = {}_{\alpha,\beta}H_2(0,1)=0$; isto é, ${}_{\alpha,\beta}H_n$ é decisiva;

(iii) ${}_{\alpha,\beta}H_n(p_1,\dots,p_n)$ é aditiva

(iv) ${}_{\alpha,\beta}H_2(1/2,1/2) = 1$

(v) ${}_{\alpha,\beta}H_n(p_1,\dots p_n) \leq {}_{\alpha,\beta}H_n(1/n,\dots,1/n)$

$= \log(n)$.

1.4.5. - ENTROPIA DE ORDEM α E TIPO β

Tal entropia foi introduzida por Shanna e Mittal e, recentemente, foi estudada por Zatelli [29].

Nessa seção tal medida de informação é definida e, como vimos fazendo nas seções anteriores, também daremos aqui uma relação de propriedades satisfeitas por esta entropia. Vale salientar, ainda, que esta medida é também conhecida por entropia paramétrica generalizada ou, ainda, por entropia de ordem α e grau β.

Definição (entropia de ordem α e tipo β)

A entropia de ordem $\alpha \neq 1$ e tipo $\beta \neq 1$ (entropia de Mittal) de uma distribuição de probabilidade $P=(p_1,\dots,p_n) \in \Gamma_n$, aqui denotada por $H_n^*(P;\alpha,\beta)$ $(n=2,3,\dots)$, é definida por

$$H_n(p_1,\dots,p_n;\alpha,\beta) = (2^{1-\beta}-1)^{-1}\left[\left(\sum_{i=1}^{n} p_i^\alpha\right)^{\frac{\beta-1}{\alpha-1}}-1\right], \tag{1.4.5}$$

onde $\alpha,\beta \in \mathbb{R}$; $\alpha,\beta >0$.

Para esta entropia, facilmente, obtemos os seguintes casos particulares:

1. $\lim\limits_{\alpha\to 1} H_n^*(p_n,\dots,p_n;\alpha,\beta) = {}_{\alpha}H_n(p_1,\dots,p_n)$ (entropia de Renyi):

2. $\lim\limits_{\alpha\to 1}\lim\limits_{\beta\to 1} H_n(p_1,\dots,p_n;\alpha,\beta) = H_n(p_1,\dots,p_n)$ (entropia de Shannon)

3. Quando $\alpha=\beta\neq 1$, $H_n(P;\alpha,\beta)$ reduz à entropia de grau α; isto é:

$$H_n^*(P;\alpha,\alpha) = (2^{1-\alpha} - 1)^{-1} \left(\sum_{i=1}^{n} p_i^{\alpha} - 1 \right), \quad \alpha > 0, \quad \alpha \neq 1.$$

4. Fazendo $\gamma = 1/\alpha$ e supondo $\beta = 2 - (1/\alpha)$, obtemos que

$$H_n^*(P;1/\gamma,2-\gamma) = {}_{\gamma}H_n(P),$$

onde:

$${}_{\gamma}H_n(P) = (2^{\gamma-1} - 1)^{-1} \left[\left(\sum_{i=1}^{n} p_i^{1/\gamma} \right)^{\gamma} - 1 \right], \quad \gamma \neq 1, \quad P \in \Gamma_n,$$

é a entropia gama generalizada.

5. Tomando $\alpha = R$ e supondo $\beta = 2-(1/R)$, obtemos:

$$H_n^*(P;R,2-(1/R)) = H_R(P) \left[\frac{1-R}{(2^{1-R/R} - 1)R} \right], \quad R > 0, \quad R \neq 1,$$

onde H_R é a informação R-Norm, a qual definiremos na próxima seção.

Algumas propriedades satisfeitas por H_n^* são:

(i) $H_n^*(p_1,\ldots,p_n;\alpha,\beta)$ é uma função simétrica de seus argumentos,

(ii) $H_n^*(p_1,\ldots,p_n;\alpha,\beta)$ é expansível,

(iii) $H_n^*(p_1,\ldots,p_n;\alpha,\beta)$ é decisiva;

(iv) $H_n^*(p_1,\ldots,p_n;\alpha,\beta)$ é não aditiva:

$$\begin{aligned} H_n^*(p_1q_1,\ldots,p_1q_m;\ldots;p_nq_1,\ldots,p_nq_m;\alpha,\beta) &= H_n^*(p_1,\ldots,p_n;\alpha,\beta)\, H_m^*(q_1,\ldots,q_m;\alpha,\beta), \\ &+ A_{\beta} H_n^*(p_1,\ldots,p_n;\alpha,\beta)\, H_m^*(q_1,\ldots,q_m;\alpha,\beta). \end{aligned}$$

Onde $A_{\beta} := (2^{1-\beta} - 1)^{-1}$.

(v) $H_n^*(p_1,\ldots,p_n;\alpha,\beta)$ é côncava se $\alpha > 0$ e $\beta \geq 2 - (1/\alpha)$.

(vi) $$\begin{aligned} H_n^*(p_1,\ldots,p_n;\alpha,\beta) &\leq H_n^*(1/n,\ldots,1/n;\alpha,\beta) \\ &= (n^{1-\beta} - 1)/(2^{1-\beta} - 1). \end{aligned}$$

(vii) $H_n^*(p_1,\ldots,p_n;\alpha,\beta) \geqq 0$, com a igualdade acontecendo se, e somente se, $P_i = 1$ para algum i e zero para todos os outros,

(viii) $H_2^*(1/2,1/2;\alpha,\beta) = 1$.

Finalizamos este parágrafo introduzindo a medida de informçaõ R-Norm, mencionado no caso particular 5, acima.

1.4.6 - MEDIDA DE INFORMAÇÃO R-NORM

A medida de informação R-Norm é introduzida e, a seguir a exemplo das demais seções, listamos algumas propriedades satisfeitas por esta medida de informação.

Definição (Medida de informação R-Norm)

A medida de informação R-Norm de uma distribuição de probabilidade $P = (P_1,\ldots p_n) \in \Gamma_n$, denotamos por $H_R(P)$, é definida por:

(1.4.6) $$H_R(p_1,\ldots,p_n) = \frac{R}{R-1}\left[1 - \left(\sum_{i=1}^{n} p_i^R\right)^{1/R}\right],$$

onde $R > 0$. $R \neq 1$, $R \in \mathbb{R}$

Como vemos, H_R é uma função real definida sobre Γ_n tomando valores em $\mathbb{R}$, onde $n = 2,3,\ldots$.

I. PROPRIEDADES ALGÉBRICAS DE H_R:

(i) $H_R(p_1,\ldots p_n)$ é uma função simétrica de $p_1,\ldots,p_n$;

(ii) $H_R(P)$ é expansível;

(iii) $H_R(P)$ é decisiva: $H_R(1,0) = H_R(0,1) = 0$;

(iv) $H_R(P)$ é recursiva;

(v) $H_R(P)$ é pseudo-aditiva; i.e, se P e Q são distribuições independentes, então:

$$H_R(P^*Q) = H_R(P) + H_R(Q) - \Big[(R-1)/R\Big]H_R(P)H_R(Q).$$

II. PROPRIEDADES ANALÍTICAS DE H_R:

(i) $H_R(P) \geqq H_R(1,0,\ldots 0) = 0$; i.e, H_R é não-negativa;

(ii) $H_R(p_1,\ldots,p_n) \leqq H_R(1/n,\ldots,1/n)$

$$= (R/R-1)\left(1 - n^{1-R/R}\right);$$

(iii) $H_R(P)$ é uma função monótona de seus argumentos;

(iv) $H_R(P)$ é uma função contínua em R,

(v) $H_R(P)$ é uma função côncava para todo $P \in \Gamma_n$.

As demonstrações das propriedades mencionadas acima podem ser vistas, por exemplo, na referência [4].

1.5. A DESIGUALDADE DE INDEPENDÊNCIA

O objetivo dessa seção é tão somente introduzir a "A desigualdade de independência" para uma sequência de funções (entropias) $J_n:\Gamma_n \to \mathbb{R}$ $(n=2,3,\ldots)$. Isto é feito na seguinte

Definição (desigualdade de independência)

Dizemos que uma sequência de funções (entropias) $J_n:\Gamma_n \to \mathbb{R}$ $(n=2,3,\ldots)$ satisfaz a desigualdade de independência se, e somente se,

$$J_{nm}(\pi_{11},\pi_{12},\ldots,\pi_{1m};\ldots;\pi_{n1},\pi_{n2},\ldots,\pi_{nm}) \tag{1.5.1}$$

$$\leqq \quad J_{nm}(p_1q_1,p_1q_2,\ldots,p_1q_m;\ldots;p_nq_1,p_nq_2,\ldots,p_nq_m),$$

para todo n, m e toda distribuição $(\pi_{11},\ldots,\pi_{1m};\ldots;\pi_{n1},\ldots,\pi_{nm})$ $\in \Gamma_{nm}$.

Em palavras, a desigualdade de independência diz que a informação obtida da combinação de dois experimentos é máxima quando os experimentos são independentes.

Em forma abreviada, a desigualdade de independência diz que

$$J_{nm}(\pi_{ij}) \leqq J_{nm}(p_i q_i) \tag{1.5.2}$$

No capítulo 2, a seguir, fazemos um estudo dessa desigualdade relacionando-a com as entropias de Shannon, Rényi, Daroczy Sharma-Taneja e, finalmente, com a entropia generalizada de ordem (α,β). Deixamos para o capítulo 3 o estudo relacionando-a com a entropia de Mittal.

CAPÍTULO II

APLICAÇÕES DE DESIQUALDADE DE INDEPENDÊNCIA

2.1. INTRODUÇÃO

Apresentamos neste capítulo, diferentes exemplos de distribuições de probabilidades e canais de comunicação, para os quais a desigualdade de independência se verifica. Para isso, consideramos diferentes medidas de informações, tais como: a entropia de Shannon, entropia de Renyi (entropia de ordem α), entropia de Daróczy (entropia de grau α), entropia de Sharna-Taneja (entnopia de grau (α,β)) e finalmente, a entropia de ordem (α,β).

Mostramos que a entropia de Shannon satisfaz a desigualdade de independência para todas as distribuições de probabilidades $P \in \Delta_n$, enquanto que as demais satisfazem a desigualdade da questão somente para certas classes de distribuições. Discutimos, também, a relação existente entre a desigualdade de independência e as propriedades de aditividade e subaditividade. Em outras palavras, o estudo que desenvolvemos neste segundo capítulo, pode ser resumido na seguinte afirmação:

A desigualdade de independência á satisfeita pelo seguinte conjunto de entropias:

(i) H_n^{α}, $\alpha > 0$, $\alpha \neq 1$

(ii) ${}_{\alpha}H_n$, $\alpha > 0$, $\alpha \neq 1$

(iii) $H_n^{\alpha\,\beta}$, $\beta \geq 1 > \alpha \geq 0$ (ou $\alpha > 1 > \beta \geq 0$)

(iv) ${}_{\alpha\,\beta}H_n$, $\beta \geq 1 > \alpha \geq 0$ (ou $\alpha > 1 > \beta \geq 0$);

Se uma das seguintes condições é satisfeita:

1. A distribuição de probabilidades de entrada p_i, $i=1,\ldots,n$ é uniforme (em partícular, se a distribuição conjunta de probabilidades π_{ij}, $i=1,\ldots,n$; $j = 1,\ldots,m$ é uniforme).

2. A distribuição de probabilidades de saída q_j, $j=1,\ldots m$, é uniforme (em particular, se a distribuição condicional de probabilidades q_{ij}, $i = 1 \ldots$; $j=1,\ldots,m$ é uniforme).

3. A matriz de probabilidade condicional $(q_{ij})_{n,m}$ tem a propriedade de permutação-linha.

E mais, a entropia H_n satisfaz a desigualdade de independência qualquer que seja a distribuição de probebilidades.

2.2. ENTROPIA DE SHANNDN E A DESIGUALDADE DE INDEPENDÊNCIA

Consideremos um canal discreto com alfabeto de entrada dado por $\{x_1,x_2,\dots,x_n\}$ e alfabeto de saída $\{y_1,y_2,\dots y_m\}$, caracterizado pela (n,m) - matriz de transição

$$\mathcal{Q} = (q_{ij}), \quad (i=1,\dots,n;\ j=1,\dots,m) \text{ oom } q_{ij} \geq 0$$

$$\sum_{j=i}^{m} q_{ij} = 1 \qquad (i=1,\dots,n).$$

Sejam X e Y variáveis aleatórias com distribuções de probabilidades $P = (p_1,p_2,\dots,p_n) \in \Gamma_n$ e $Q = (q_1,q_2,\dots,q_m) \in \Gamma m$, respectivamente. Então podemos considerar os alfabetos de entrada e de saída como os alfabetos de valores para X e Y, respectivamente, onde as distribuições de probabilidades para X e Y são dadas por

$$P(X=x_i) = p_1 \quad (i=1,\dots,n)$$

$$P(Y=y_j) = q_j = \sum_{i=1}^{n} p_i q_{ij} \quad (j=1,\dots,m)$$

$$P(Y=j_i \ / \ X=j_i) = q_{ij} \ (i=1,\dots,n;\ j=1,\dots,m)$$

$$P(X=y_i,\ Y=_{ji}) = \P_{ij} \ (i=1,\dots,n;\ j=i,\dots,m)$$

como no capítulo I, seção 1.2.

Os resultados seguintes enfatizam a importância da desigualdade de independência, ao mesmo tempo em que dão diferentes exemplos de distribuições de probabilidades e canais de comunicações, para os quais esta desigualdade se verifica.

Teorema 2.2.1

Seja $J_n:\Gamma_n \to \mathbb{R}$ (n=2,3,...) uma sequência aditiva de funções, definidas sobre Γ_n tomando valores no conjunto $\mathbb{R}$ dos números reais. Então J_n é sub-aditiva se, e somente se, satisfaz a desigualdade de independência.

Demonstração: A propriedade de aditividade, para uma sequência de funções $J_n:\Gamma_n \to \mathbb{R}$(n=2,3,...), pode ser escrita na forma abreviada.

(2.2.1) $$J_{nm}(p_i q_j) = J_n(p_i) + J_m(q_j),$$

enquanto a propriedade de subatividade pode, também na forma abreviada ser dada por

(2.2.2) $$J_{nm}(\pi_{ij}) \leqq J_n(p_i) + J_m(q_j),$$

uma vez que

$$p_i = \sum_{j=1}^{m} \pi_{ij} \quad \text{e} \quad q_j = \sum_{i=1}^{n} \pi_{ij}$$

Assim, se $J_n:\Gamma_n \to \mathbb{R}$ (n=2,3,...) é uma sequência aditiva de funções, sendo esta também subaditiva, combinando as relações (2.2.1) e (2.2.2), obtemos

$$J_{nm}(\pi_{ij}) \leqq J_{nm}(p_i q_i),$$

a qual é a forma abreviada da desigualdade de independência dada por (1.5.2) e, portanto, se uma sequência aditiva de funções $J_n:\Gamma_n \to \mathbb{R}$ (n=2,3,...) é também subaditiva, esta satisfaz a desigualdade de indepdência.

A recíproca pode ser vista facilmente. Para isto, basta combinar as relações (1.5.2) e (2.2.1).

Corolário 2.2.1

A entropia de Shannon, H_n, satisfaz a desigualdade de independência qualquer que seja a distribuição de probabilidades.

Demonstração: De fato, como vimos (seção (1.3)) a entropia de Shannon é aditiva e subaditiva e, portanto, pelo teorema acima segue-se o resultado, o que prova a nossa afirmação.

Pelo teorema 2.2.1 observamos, ainda, que se uma entropia satisfaz a desigualdade de independência e é aditiva para certas distribuições de probabilidades, então para estas mesmas distribuições, esta entropia também será subaditiva.

2.3. ENTROPIA DE GRAU α E A DESIGUALDADE DE INDEPENDÊNCIA

Na proposição seguinte damos um exemplo no qual mostramos que a desigualdade de independência não é satisfeita pela entropia H_n^α (entropia de grau α).

Proposição 2.3.1

Existe α e distribuições de probabilidades $P = (p_1, p_2, \ldots p_m) \in \Gamma_n$, tal que a entropia $H_n^\alpha(P)$ não satisfaz a desigualdade de independência.

Demonstração: Sejam $\alpha = 2$, $n=2=m$, $p_1=0.1$, $p_2 = 0.9$. Consideremos, também, a seguinte (2,2) - matriz

$$(\pi_{ij}) = \begin{pmatrix} 0.04 & 0.09 \\ 0.34 & 0.53 \end{pmatrix}$$

Então,

$$q_1 = \sum_{i=1}^{2} \pi_{ij} = 0.38,$$

$$q_2 = \sum_{i=1}^{2} \pi_{ij} = 0.62 \quad (= 1-q_1),$$

$$\sum_{i,j} \pi_{ij}^\alpha = \sum_{i,j} \pi_{ij}^2 = 0.4062 \tag{2.3.1}$$

e, finalmente,

$$\sum_{i,j} (p_i q_j)^\alpha = \sum_{i,j} (p_i q_j)^2 = 0.4336 . \tag{2.3.2}$$

Para $\alpha = 2$, $A_\alpha := (2^{1-\alpha}-1)^{-1} = -2 < 0$ e, portanto, das relações (2.3.1) e (2.3.2), obtemos que

$$H^\alpha_{nm}(\pi_n,\ldots,\pi_{nm}) > H^\alpha_{nm}(p_1q_1,\ldots,p_nq_m)$$

Assim vemos que a desigualdade de independência não é satisfeita.

Mostraremos, agora, que para certas classes de distribuições de probabilidades $P = (p_1 p_2,\ldots,p_n) \in \Gamma_n$, a entropia $H^\alpha_n(P)$ $(\alpha > 0),$ $\alpha \neq 1)$ satisfaz a desigualdade de independência. Para isto, consideramos a seguinte (n,n) - matriz $Q = (q_{ij})$, definida por

$$q_{ij} = \begin{cases} 1-p, & \text{se } i=j \\ \dfrac{p}{n-1}, & \text{se } i\neq j \end{cases}$$

onde

$$0 \leq p \leq 1$$

Seja $\qquad p_i = 1/n,\ i=1,\ldots,n$

É fácil ver que

$$\pi_{ij} = \frac{1}{n} q_{ij},$$

$$q_j = \frac{1}{n},$$

$$p_i q_j = \frac{1}{n^2}, \quad (i=1,\ldots,n;\quad j=1,\ldots,n)$$

Para H^α_n, a desigualdade de independência tem a forma

$$(2^{1-\alpha}-1)^{-1}\left(\sum_{i,j} \pi^\alpha_{ij} - 1\right) \leq (2^{1-\alpha}-1)^{-1}\left(\sum_{i,j} (p_i q_j)^\alpha - 1\right);$$

isto é,

$$(2.3.3) \qquad A_\alpha \sum_{i,j} \pi^\alpha_{ij} \leq A_\alpha \sum_{i,j} (p_i q_j)^\alpha$$

onde

$$A_\alpha := (2^{-1} -1)^{-1}$$

é fácil ver, ainda, que

$$\sum_{i,j} (p_i q_j)^\alpha = (n^{1-\alpha})^2$$

e

$$\sum_{i,j} \P_{ij}^\alpha = n^{1-\alpha}\Big((1-p)^\alpha + p^\alpha (n-1)^{1-\alpha}\Big).$$

Seja

$$f(p) := (1-p)^\alpha + p^\alpha (n-1)^{1-\alpha}$$

Então,

$$f'(p) = -\alpha(1-p)^{\alpha-1} + \alpha p^{\alpha-1} . (n-1)^{1-\alpha} ,$$

$$f''(p) = \alpha(\alpha-1)\Big((1-p)^{\alpha-2} + (n-1)^{1-\alpha} . p^{\alpha-2} \Big),$$

$f'(p) = 0$ se, e somente se, $p = \frac{n-1}{n}$, i.e, f tem um ponto crítico em $p_c = (n-1)/n$ e $f(p_c) = n^{1-\alpha}$

Temos dois casos a considerar:

(i) $0 \leqq \alpha < 1$:

Neste caso, $A_\alpha > 0$ e p_c é um ponto de máximo. Daí, segue-se que a desigualdade (2.3.3), a qual é equivalente a

$$A_\alpha . n^{1-\alpha} . f(p) \leqq A_\alpha . (n^{\alpha-1})^2,$$

é satisfeita.

(ii) $\alpha > 1$:

Aqui, $A_\alpha < 0$ e p_c um ponto de mínimo e, novamente, a desigualdade (2.3.3) é satisfeita

Definição 2.3.1

Uma função real e diferenciável Ψ é dita ser "côncava" em um intervalo (a,b), se esta é definida e possui derivada segunda em (a,b) e satisfaz

$$\Psi''(x) \leq 0,$$

para todo $x \in (a,b)$

Uma tal função satisfaz, ainda, a desigualdade

$$(2.3.4) \qquad \sum_{i=1}^{n} p_i \Psi(x_i) \leq \Psi\left(\sum_{i=1}^{n} p_i x_i \right),$$

para todo $x_i \in (a,b)$; $i=1,\ldots,n$; $\sum_{i=1}^{n} p_i = 1$ (ver referência [1]).

Se $\Psi(x) \leq 0$, para todo $x \in (a,b)$, então Ψ é dita ser uma função "convexa" e, neste caso, a desigualdade (2.3.4) fica invertida

Proposição 2.3.2

A desigualdade de independência é satisfeita pela entropia H_n^{α} ($\alpha \geq 0$, $\alpha \neq 1$) para todas distribuições de probabilidades $P = (p_1, p_2 \ldots p_n) \in \Gamma_n$, que dão origem a alfabetos de saída com distribuições de probabilidades equiprováveis, i.é, $Q = (q_1, q_2, \ldots, q_n) \in \Gamma_m$ tal que $q_j \neq 1/m$, $j = 1,\ldots,m$.

Demonstração: A prova da afirmação acima consiste de duas partes, quais sejam:

(i) $0 \leq \alpha < 1$:

Esta variação de α nos dá que $A\alpha = (2^{1-\alpha}-1)^{-1} > 0$

Usando o fato de que a função $x \quad x^{\alpha}$ ($0 \leq \alpha < 1$) é côncava, obtemos que

$$\sum_j q_{ij}^{\alpha} \leq \sum_j q_j^{\alpha}$$

para todo $i=1,\ldots,n$. Multiplicado por p_i e somando sobre i, temos que

$$\sum_{i,j} (p_i q_{ij})^\alpha \leq \sum_{i,j} (p_i q_j)^\alpha$$

ou,

$$\sum_{i,j} \pi_{ij}^\alpha - 1 \leq \sum_{i,j} (p_i q_i)^\alpha - 1$$

Multiplicando pelo fator A_α, o qual é positivo, obtemos o resultado.

(ii) $\alpha > 1$:

Neste caso, a função $x \longmapsto x^\alpha$ é convexa e, temos

$$\sum_j q_{ij}^\alpha \geq \sum_j q_j^\alpha,$$

que nos dá

$$\sum_{i,j} (p_i q_{ij})^\alpha \geq \sum_{i,j} (p_i q_j)^\alpha$$

Como, neste caso, $A_\alpha < 0$, obtemos

$$A_\alpha \cdot \left(\sum_{i,j} \pi_{ij}^\alpha - 1 \right) \leq A\alpha \left(\sum_{i,j} (p_i q_j)^\alpha - 1 \right)$$

e novamente, a desigualdade de independência se verifica por H_n^α.

Proposição 2.3.3

Seja $(q_{ij})_{nm}$ a (n,m) - matriz de probabilidade condicional, cujas linhas são as permutações de um mesmo conjunto de m números. Então a entropia H_n^α $(\alpha \leq 0,\ \alpha \neq 1)$ satisfaz a desigualdade de independência.

Demonstração: Seja $(a_1, a_2, \ldots, a_m)$ um tal conjunto de n números, e seja $\Sigma_{\ell=1}^m a_\ell^\alpha = A$. Então,

$$\sum_j q_{ij}^{\alpha} = A, \quad (i=1,\dots,n),$$

$$\sum_j \sum_j p_i q_{ij}^{\alpha} = A;$$

donde,

$$\sum_j q_{ij}^{\alpha} = \sum_j \sum_i p_i q_{ij}^{\alpha}, \quad (i=1,\dots,n)$$

Para $0 \leqq \alpha < 1$, temos que

$$\sum_j q_{ij}^{\alpha} \leq \sum_j q_j^{\alpha}$$

A partir daqui, a demonstração é análoga à da proposição anterior.

Na proposição seguinte consideramos o caso de um canal com alfabeto de entrada tendo distribuição de probabilidades uniforme e, como veremos, a desigualdade em questão é, novamente, satisfeita por $\bar{H}_n^{\alpha}$.

Proposição 2.3.4

Se a distribuição de probabilidades de entropia (p_i) é uniforme, então a desigualdade de independência é satisfeita pela entropia H_n^{α}.

Demonstração: Sejam

$$L := \sum_{i,j} (p_i q_{ij})^{\alpha} = \left(\frac{1}{n}\right)^{\alpha} \cdot \sum_{i,j} q_{ij}^{\alpha},$$

$$R := \sum_{i,j} (p_i q_j)^{\alpha} = \left(\frac{1}{n}\right)^{\alpha} \cdot n^{1-\alpha} \cdot \sum_j \left(\sum_i q_{ij}\right)^{\alpha}.$$

Para $0 \leqq \alpha < 1$, facilmente, vemos que $L \leqq R$. Como neste caso $A_{\alpha} := (2^{1-\alpha}-1)^{-1}$ é positivo, segue-se o resultado desejado. Para $\alpha > 1$, temos que A_{α} é negativo e $L \geqq R$ e, novamente, a desigualdade em questão é satisfeita. Dessa forma a prova está completa.

Corolário 2.3.1

A desigualdade de independência é satisfeita por H_n^α se a distribuição conjunta de probabilidades é uniforme.

Demonstração: De fato, sendo a distribuição conjunta uniforme, temos que $p_i = 1/n$ para todo i. Assim, temos o resultado.

Corolário 2.3.2

A desigualdade de independência é satisfeita por H_n^α se a distribuição condicional de probabilidades é uniforme.

Demonstração: Neste caso a propriedade de permutação de linhas da matriz $(q_{ij})_{n,m}$ é satisfeita. Isto completa a demonstração.

2.4. ENTROPIA DE RÉNYI E A DESIGUALDADE DE INDEPENDÊNCIA

No lema a seguir mostraremos que uma entropia não tem que ser aditiva ou subaditiva para satisfazer a desigualdade de independência.

Lema 2.4.1

Seja $\alpha \in R$, $\alpha \neq 1$. Então, para uma dada distribuição de probabilidades $P \in \Gamma_n$, se H_n^α (P) satisfaz a desigualdade de independência, temos que ${}_\alpha H_n$ (P) (entropia de Rényi) também verifica esta desigualdade, e vice-versa.

Demonstração: A desigualdade de independência para H_n^α, como já vimos, tem a forma

$$(2.4.1) \qquad A_\alpha \left(\sum_{i,j} \pi_{ij}^{\alpha} - 1 \right) \leq A_\alpha \left(\sum_{i,j} (p_i q_j)^{\alpha} - 1 \right),$$

isto é,

$$(2.4.2) \qquad A_\alpha \sum_{i,j} \pi_{ij}^{\alpha} \leq A_\alpha \sum_{i,j} (p_i q_j)^{\alpha} ,$$

onde $\qquad A_\alpha = (2^{1-\alpha} - 1)^{-1}.$

Para ${}_\alpha H_n$, a desigualdade de independência, toma a forma

$$(2.4.3) \qquad \overline{A}_\alpha \ \log \left(\sum_{i,j} \pi_{ij}^\alpha \right) \leq \overline{A}_\alpha \ \log \left(\sum_{i,j} (p_i q_j)^\alpha \right),$$

onde $\overline{A}_\alpha := 1/(1-\alpha)$.

Evidentemente que (2.4.2) e (2.4.3) são equivalentes (i.e, uma implica na outra) pois $x \longmapsto \log(x)$ é uma função crescente, A_α e $\overline{A}_\alpha$ têm o mesmo sinal para cada valor do parâmetro α.

Com efeito, a relação entre $H_n^\alpha(P)$ e ${}_\alpha H_n(P)$, $P \in \Gamma_n$, pode ser expressa na forma

$${}_\alpha H_n(P) = \overline{A}_\alpha \log(1 + (1/A_\alpha) \cdot H_n^\alpha(P)),$$

isto é,

$${}_\alpha H_n(P) = F(H_n^\alpha(P)),$$

onde

$$F(x) = \overline{A}_\alpha \cdot \log(1 + (x/A_\alpha)),$$

a qual é uma função crescente de x. Isto completa a demonstração do lema.

Na seção 1.4, vimos que a entropia H_n^α era subaditiva, enquanto que a entropia ${}_\alpha H_n$ não.

Segue do lema acima que todas as afirmações vistas no caso da entropia H_n^α valem, igualmente, para a entropia ${}_\alpha H_n$.

2.5. ENTROPIA DE SHARMA-TANEJA E A DESIGUALDADE DE INDEPENDÊNCIA

Nesta seção, mostraremos que as proposições vistas no parágrafo 2.3, valem, igualmente, no caso da entropia generalizada de grau (α,β) (entropia de Sharma-Taneja). Temos, portanto, a seguinte:

Proposição 2.5.1

Existem α,β e distribuições de probabilidades $P \in \Gamma_n$, tais que a entropia $H_n^{\alpha,\beta}(P)$ não satisfaz a desigualdade de independência.

Demonstração: Para $H_n^{\alpha,\beta}$, a desigualdade de independência se apresenta como na forma seguinte:

$$(2.5.1) \qquad \frac{1}{2^{1-\alpha}-2^{1-\beta}} \cdot \sum_{i,j} (\pi_{ij}^{\alpha} - \pi_{ij}^{\beta}) \leq \frac{1}{2^{1-\alpha}-2^{1-\beta}} \cdot \sum_{i,j} ((p_i q_j)^{\alpha} - (p_i q_j)^{\beta}).$$

Considerando $\alpha = 2$, $\beta = 1.5$ e usando a distribuição de probabilidades dada na demonstração da proposição 2.3.1, juntamente com a matriz $(\pi_{ij})_{2,2}$ que lá explicitamos, obtemos que

$$\sum_{i,j} \pi_{ij}^{\alpha} = 0.4062, \qquad \sum_{i,j} (p_i q_j)^{\alpha} = 0.4336,$$

enquanto que

$$\sum_{i,j} \pi_{ij}^{\beta} = 0.6190, \qquad \sum_{i,j} (p_i q_j)^{\beta} = 0.6403.$$

Daí, temos que

$$\sum_{i,j} (\pi_{ij}^{\alpha} - \pi_{ij}^{\beta}) = -0.2128$$

e

$$\sum_{i,j} ((p_i q_j)^{\alpha} - (p_i q_j)^{\beta}) = -0.2067.$$

Para os valores dos parâmetros α e β, considerados acima, temos que

$$\frac{1}{2^{1-\alpha} - 2^{1-\beta}} = -2(1+\sqrt{2}),$$

o qual é negativo, e portanto, a desigualdade da independência não se verifica, como queríamos.

O análogo da proposição 2.3.2, no caso da entropia $H_n^{\alpha,\beta}$ é a seguinte:

Proposição 2.5.2

A desigualdade de independência é satisfeita pela entropia $H_n^{\alpha,\beta}$ $(0 \leq \alpha < 1 < \beta$ ou $0 \leq \beta < 1 < \alpha)$ para as distribuições de probabilidades que dão origem a distribuições de probabilidades de saída equiprováveis $(q_j = 1/m;\ j = 1, \ldots, m)$.

Demonstração: Já vimos que para $0 \leq \alpha < 1$, ocorre

$$\sum_{i,j} \pi_{ij}^{\alpha} \leq \sum_{i,j} (p_i q_j)^{\alpha}. \tag{2.5.2}$$

Para $\beta > 1$, a função $x \longmapsto x^{\beta}$ é convexa e, assim, obtemos

$$\sum_{i,j} \pi_{ij}^{\beta} \geq \sum_{i,j} (p_i q_j)^{\beta}. \tag{2.5.3}$$

Subtraindo (2.5.3) de (2.5.2), segue-se que

$$\sum_{i,j} (\pi_{ij}^{\alpha} - \pi_{ij}^{\beta}) \leq \sum_{i,j} \left((p_i q_j)^{\alpha} - (p_i q_j)^{\beta} \right).$$

Multiplicando a desigualdade acima pelo fator $A_{\alpha,\beta} := (2^{1-\alpha} - 2^{1-\beta})^{-1}$, o qual é positivo, obtemos

$$A_{\alpha,\beta} \sum_{i,j} (\pi_{ij}^{\alpha} - \pi_{ij}^{\beta}) \leq A_{\alpha,\beta} \sum_{i,j} \left((p_i q_j)^{\alpha} - (p_i q_j)^{\beta} \right),$$

a qual é a desigualdade de independência para $H_n^{\alpha,\beta}$.

Proposição 2.5.3

Seja $(q_{ij})_{n,m}$ a (n,m) - matriz cujas linhas são as permutações de um mesmo conjunto de m números. Então a entropia $H_n^{\alpha,\beta}$ $(0 \leq \alpha < 1 < \beta$ ou $0 \leq \beta < 1 < \alpha)$ satisfaz a desigualdade de independência.

Demonstração: Seja $(a_1, a_2, \ldots, a_m)$ um tal conjunto de m números, e seja $\sum_{\ell=1}^{m} a_\ell = A$. Para $0 \leq \alpha < 1$, temos que

$$\sum_j q_{ij}^{\alpha} \leq \sum_j q_j^{\alpha},$$

como anteriormente. Para $\beta > 1$, concluimos que

$$\sum_j q_{ij}^{\beta} \geq \sum_j q_j^{\beta},$$

a qual implica em

$$\sum_{i,j} \pi_{ij}^{\beta} \geq \sum_{i,j} (p_i q_j)^{\beta}.$$

Dessa forma, a prova da afirmação acima está completa pelo que fizemos no caso anterior.

Consideramos na proposição seguinte, o caso de um canal com alfabeto de entrada tendo distribuição uniforme de probabilidades.

Proposição 2.5.4

A desigualdade de independência é satisfeita pela entropia $H_n^{\alpha,\beta}$ $(0 \leq \alpha < 1 \leq \beta$ ou $0 \leq \beta < 1 < \alpha)$, se a distribuição de probabilidades de entrada é uniforme.

Demonstração: A prova neste caso é similar à da proposição 2.5.2 e, por isso, a omitiremos.

Corolário 2.5.1

A desigualdade de independência é satisfeita por $H_n^{\alpha,\beta}$, se a distribuição conjunta de probabilidades é uniforme.

Demonstração: Similar à do Corolário 2.3.1.

A afirmação seguinte pode ser considerada como um corolário para uma das proposições (2.5.2) e (2.5.3).

Corolário 2.5.2.

A desigualdade de independência é satisfeita por $H_n^{\alpha,\beta}$, se a distribuição condicional de probabilidades é uniforme.

Demonstração: A prova desse corolário é similar à que foi dada ao corolário 2.3.2. Por isso, a omitiremos.

2.6. ENTROPIA DE ORDEM (α,β) E A DESIGUALDADE DE INDEPENDÊNCIA

A desigualdade de independência, no caso da entropia ${}_{\alpha,\beta}H_n$ (entropia de ordem (α,β), é equivalente a

$$(2.6.1)\qquad \frac{1}{\beta-\alpha}\cdot\log\left(\frac{\sum\limits_{i,j}\pi_{ij}^{\alpha}}{\sum\limits_{i,j}\pi_{ij}^{\beta}}\right)\leq\frac{1}{\beta-\alpha}\cdot\log\left(\frac{\sum\limits_{i,j}(p_iq_j)^{\alpha}}{\sum\limits_{i,j}(p_iq_j)^{\beta}}\right),$$

a qual implica na desigualdade

$$(2.6.2)\qquad \frac{\sum\limits_{i,j}\pi_{ij}^{\alpha}}{\sum\limits_{i,j}\pi_{ij}^{\beta}}\geq\frac{\sum\limits_{i,j}(p_iq_j)^{\alpha}}{\sum\limits_{i,j}(p_iq_j)^{\beta}}$$

para $\alpha > \beta$ e, na desigualdade

$$(2.6.3)\qquad \frac{\sum\limits_{i,j}\pi_{ij}^{\alpha}}{\sum\limits_{i,j}\pi_{ij}^{\beta}}\leq\frac{\sum\limits_{i,j}(p_iq_j)^{\alpha}}{\sum\limits_{i,j}(p_iq_j)^{\beta}}$$

se $\alpha < \beta$.

É fácil ver que o conjunto de afirmações feitas na seção anterior vale, igualmente, no caso da entropia de ordem (α,β), ${}_{\alpha,\beta}H_n$. Dado a semelhança das demonstrações as omitiremos aqui.

Portanto, podemos ainda ter o seguinte

Corolário 2.6.1

As entropias aditivas

(i) ${}_{\alpha}H_n$, $\alpha \geq 0$, $\alpha \neq 1$.

(ii) ${}_{\alpha\beta}H_n$, $0 \leq \alpha < 1 < \beta$

São subaditivas para algumas das distribuições mencionadas nas proposições citadas.

Demonstração: Já vimos que ${}_{\alpha}H_n$ e ${}_{\alpha,\beta}H_n$ são aditivas. Assim, para algumas das distribuições mencionadas nas proposições citadas, as entropias (i) e (ii) são aditivas e satisfazem a desigualdade de independência, e portanto, pelo teorema 2.2.1 estas são, também, subaditivas.

CAPÍTULO III

A DESIGUALDADE DE INDEPENDÊNCIA PARA A ENTROPIA DE MITTAL

3.1. INTRODUÇÃO

Aborda-se, neste capítulo, o problema que consiste em mostrar a desigualdade de independência no caso da entropia de ordem α e tipo β (entropia de Mittal) e, como já tivemos a oportunidade de frizar, este resultado constitui basicamente a parte original do nosso trabalho.

A solução aqui apresentada para este problema, é particularizada para certas classes de distribuições de probabilidades e para certos valores dos parâmetros α e β que aparecem na definição da entropia em questão. Para isto, usaremos as definições e notações introduzidas nos capítulos anteriores.

3.2. A DESIGUALDADE DE INDEPENDÊNCIA

Já vimos que a entropia de ordem α e tipo β de uma distribuição de probabilidades $P=(p_1, p_2, \ldots, p_n) \in \Gamma_n$ de uma variável aleatória $X = \{x_1, x_2, \ldots x_n\}$, é definida da maneira seguinte:

$$(3\text{-}2\text{-}1) \qquad H_n^*(P;\alpha,\beta) = H_n^*(p_1, p_2, \ldots, p_n; \alpha,\beta) = (2^{1-\beta}-1)^{-1}\left[\left(\sum_{i=1}^{n} p_i^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right],$$

onde $\alpha, \beta > 0$, $\alpha, \beta \neq 1$.

De (1.5.1), juntamente com (3.2.1), vemos que a desigualdade de independência, no caso da entropia de ordem α e tipo β, tem a forma dada a seguir,

$$(3.2.2) \qquad A_\beta\left[\left(\sum_{i,j} \pi_{ij}^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right] \leq A_\beta\left[\left(\sum_{i,j} (p_i q_j)^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right]$$

onde $A_\beta = (2^{1-\beta} - 1)^{-1}$

Na proposição seguinte damos um exemplo, no qual mostramos que a desigualdade de independência não se verifica no caso da entropia de Mittal, para todos os valores dos parâmetros α e β.

Proposição 3.2.1

Existem números α, β e distribuições de probabilidades $P = (p_1, p_2, \ldots, p_n)$ e Γ_n, tais que a entropia $H_n^*(p_1, p_2, \ldots, p_n)$ não satisfaz a desigualdade de independência.

Demonstração: Para provar a afirmação acima, é suficiente darmos um contra-exemplo para a desigualdade em questão. Com efeito, sejam $\alpha = 2$, $\beta = 3$, $m = 2 = n$, $p_1 = 0.1$ e $p_2 = 0.9$. Consideremos, ainda, a matriz $(\pi_{ij})_{2,2}$ dada por:

$$(\pi_{ij}) = \begin{pmatrix} 0.01 & 0.09 \\ 0.16 & 0.74 \end{pmatrix}$$

Daí, obtemos que:

$$q_1 = \sum_{i=1}^{2} \pi_{i1} = 0.01 + 0.16 = 0.17,$$

$$q_2 = \sum_{i=1}^{2} \pi_{i2} = 0.09 + 0.74 = 0.83 \; (= 1 - q_1),$$

$$\sum_{i,j} \pi_{ij}^{\alpha} = (0.01)^2 + (0.09)^2 + (0.16)^2 + (0.74)^2 = 0.5814,$$

$$\sum_{i,j} (p_i q_j)^{\alpha} = \sum_i p_i^2 \sum_j q_j^2 = (0.1)^2 + (0.9)^2 \quad (0.17)^2 + (0.83)^2 = 0.5886$$

Já que $\beta = 3$, temos que $A_\beta = -4/3$. Temos, também,

$$\left(\sum_{i,j} \pi_{ij}^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} = \left(\sum_{i,j} \pi_{ij}^{2}\right)^{2} = (0.5814)^2 = 0.3380 \tag{3.2.3}$$

e

$$\left(\sum_{i,j} (p_i q_j)^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} = \left(\sum_{i,j} (p_i q_j)^{2}\right)^{2} = (0.5886)^2 = 0.3464 \tag{3.2.4}$$

Em virtude de A_β ser negativo e de (3.2.3) ser menor do que (3.2.4), segue-se que a desigualdade (3.2.2) não se verifica, o que prova a proposição acima.

Nosso próximo passo consiste em mostrar que para certas classes de distribuições de probabilidades, a desigualdade de independência é satisfeita pela entropia de Mittal. Com efeito, consideremos a seguinte $n \times n$ matriz de transição $(q_{ij})_{n,n}$, dada por:

$$\begin{pmatrix} 1-p & \frac{p}{n-1} & \frac{p}{n-1} & \cdots & \frac{p}{n-1} \\ \frac{p}{n-1} & 1-p & \frac{p}{n-1} & \cdots & \frac{p}{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{p}{n-1} & \frac{p}{n-1} & \frac{p}{n-1} & \cdots & 1-p \end{pmatrix}_{(n,\ n)}$$

onde $0 \leq p \leq 1$. Seja, ainda, $p_i = 1/n$, para todo $i = 1, \ldots n$. Então

$$\pi_{ij} = p_i \, q_{ij} = \frac{1}{n} \, q_{ij},$$

$$q_j = \sum_{i=1}^{n} \pi_{ij} = \frac{1}{n} \sum_{i=1}^{n} q_{ij} = \frac{1}{n},$$

uma vez que $\sum_{i=1}^{n} q_{ij} = 1$, e portanto, segue que

$$p_i q_j = \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2},$$

$i = 1, \ldots, n$; $j = 1, \ldots, n$. Daí, obtemos que

$$\left[\sum_{i,j} (p_i q_j)^\alpha \right]^{\frac{\beta-1}{\alpha-1}} = \left[\sum_{i,j} \frac{1}{n^2}^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left[n^2 \cdot \left(\frac{1}{n^2} \right)^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= n^{2(1-\beta)} ,$$

$$\left[\sum_{i,j} \P^{\alpha}_{ij} \right]^{\frac{\beta-1}{\alpha-1}} = \left[\sum_{i,j} \left(\frac{1}{n} \cdot q_{ij} \right)^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left[\left(\frac{1}{n} \right) \cdot \sum_{i,j} (q_{ij})^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left[\frac{1}{n^{\alpha}} \cdot (n (1-p)^{\alpha} + (n^2 - n)) \cdot \left(\frac{p}{n-1} \right)^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left[\frac{1}{n^{\alpha}} \cdot (n (1-p)^{\alpha} + (n - 1)) \cdot \left(\frac{p}{n-1} \right)^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left[n^{1-\beta} \cdot ((1-p)^{\alpha} + p^{\alpha} \cdot (n - 1)^{1-\alpha}) \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left[n^{1-\beta} \cdot ((1-p)^{\alpha} + p^{\alpha} \cdot (n-- 1)^{1-\alpha}) \right]^{\frac{\beta-1}{\alpha-1}}$$

Seja

$$f(p) = \left((1-p)^{\alpha} + p^{\alpha}\,(n-1)^{1-\alpha}\right)^{\frac{\beta-1}{\alpha-1}}$$

Então, f tem derivada f' dada por

$$f'(p) = \frac{\beta-1}{\alpha-1}\cdot\left((1-p)^{\alpha} + p^{\alpha}(n-1)^{1-\alpha}\right)^{\frac{\beta-\alpha}{\alpha-1}}\left(-\alpha\,(1-p)^{\alpha-1} + \alpha p^{\alpha-1}.(n-1)^{1-\alpha}\right)$$

e derivada segunda, f", dada por

$$f''(p) = \frac{\beta-1}{\alpha-1}\left[\frac{\beta-\alpha}{\alpha-1}\left((1-p)^{\alpha} + p^{\alpha}(n-1)^{1-\alpha}\right)^{\frac{\beta-2\alpha+1}{\alpha-1}}\left(-\alpha(1-p)^{\alpha-1} + \alpha p^{\alpha-1}(n-1)^{1-\alpha}\right)^{2} + \right.$$

$$\left. + \left((1-p)^{\alpha} + p^{\alpha}(n-1)^{1-\alpha}\right)^{\frac{\beta-\alpha}{\alpha-1}}\cdot\left(\alpha(\alpha-1)(1-p)^{\alpha-2} + \alpha(\alpha-1)p^{\alpha-2}.\,(n-1)^{1-\alpha}\right)\right].$$

Afirmamos que f tem um ponto crítico em $p_c = (n-1)/n$. De fato, $f'(p) = 0$ se, e somente se,

$$-\alpha\,(1-p)^{\alpha-1} + \alpha p^{\alpha-1}\,(n-1)^{1-\alpha} = 0$$

$$\iff \left(\frac{p}{1-p}\right)^{\alpha-1} = (n-1)^{\alpha-1}$$

$$\iff \frac{p}{1+p} = n-1$$

$$\iff p = \frac{n-1}{n},$$

e portanto, f(p) tem um ponto crítico em $p_c = (n-1)/n$, como afirmamos. E mais,

$$f(p_c) = n^{1-\beta}.$$

Consideremos, agora, os seguintes casos:

(i) $0 \leq \alpha < 1$ e $\beta < 1$, porém $\alpha < \beta$.

Assim, $A_\beta > 0$ e $f''(p_c) < 0$. Portanto, p_c é um ponto de máximo. Dessa forma, a desigualdade (3.2.2), a qual é equivalente a

$$A_\beta \cdot n^{1-\beta} \quad f(p) \leq A_\beta \cdot (n^{1-\beta})^2,$$

é satisfeita.

(ii) $0 \leq \alpha < 1$ e $\beta < 1$, sendo que $\beta < \alpha$.

Aqui, como no caso anterior, $A_\beta > 0$ e $f''(p_c) < 0$. Logo, a desigualdade (3.2.2) é novamente satisfeita.

(iii) $0 \leq \alpha < 1$ e $\beta > 1$.

Neste caso, temos que $A_\beta < 0$ e p_c um ponto de mínimo, já que $f''(p_c) > 0$. Novamente a desigualdade (3.2.2) é satisfeita.

(iv) $0 \leq \beta < 1$ e $\alpha \geq 1$.

Evidentemente que $A_\beta > 0$ e p_c um ponto de máximo. Facilmente vemos que (3.2.2) é novamente satisfeita.

(v) $\beta \geq 1$, $\alpha \geq 1$.

Aqui $f''(p_c) > 0$, enquanto que $A_\beta < 0$. Assim, p_c é um ponto de mínimo. Novamente (3.2.2) é satisfeita.

O exemplo acima representa um canal no qual as distribuições de entrada e saída são ambas uniformes. Na próxima proposição, mostraremos a desigualdade de independência para todos os canais que possuem distribuições de probabilidades de saída uniforme. Isto significa dizer que as colunas da matriz de transição são permutações de um mesmo conjunto de n números.

Proposição 3.2.2

A desigualdade de independência é satisfeita pela entropia de Mittal para todas as distribuições de probabilidades as quais dão origem a uma distribuição de saída equiprovável ($q_j = 1/m$, $j = 1, \dots m$).

Demonstração: Temos de considerar os seguintes casos:

(i) $0 \leq \alpha < 1$, $0 \leq \beta < 1$.

Assim, $A_\beta = (2^{1-\beta} - 1)^{-1}$ é um fator positivo. Usando a concavidade da função $x \longmapsto x^\alpha$, $0 \leq \alpha < 1$, temos que:

$$\sum_j \frac{1}{m} q_{ij}^\alpha \leq \left(\sum_j \frac{1}{m} q_{ij} \right)^\alpha,$$

para todo i = 1, ..., n. Já que $\sum_j q_{ij} = 1$, segue que

$$\sum_j q_{ij}^\alpha \leq m.\left(\frac{1}{m}\right)^\alpha = \sum_j q_j^\alpha \ , \quad \text{para todo } i.$$

Multiplicando por p_i^α e somando sobre i, temos

$$\sum_i p_i^\alpha \sum_j q_{ij}^\alpha \leq \sum_i p_i^\alpha \sum_j q_j^\alpha \ ,$$

isto é,

$$\sum_{i,j} (p_i q_{ij})^\alpha \leq \sum_{i,j} (p_i q_j)^\alpha \ ,$$

ou

$$\sum_{i,j} \pi_{ij}^\alpha \leq \sum_{i,j} (p_i q_j)^\alpha \ .$$

Elevando ambos os membros da desigualdade acima à potência $(\beta - 1)/(\alpha - 1)$, a qual é positiva, depois subtraindo 1 e, em seguida, multiplicando pelo fator $A_\beta > 0$, obtemos:

$$A_\beta\left[\left(\sum_{i,j}\pi_{ij}^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right] \leq A_\beta\left[\left(\sum_{i,j}(p_i q_j)^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right],$$

a qual é a desigualdade de independência para H_n^*.

(ii) $0 \leq \alpha < 1, \ \beta > 1$.

Novamente obtemos que

$$\sum_{i,j}\pi_{ij}^{\alpha} \leq \sum_{i,j}(p_i q_j)^{\alpha}$$

Por outro lado, a potência $(\beta-1)/(\alpha-1)$ é, agora, negativa. Isto nos dá

$$\left(\sum_{i,j}\pi_{ij}^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} \geq \left(\sum_{i,j}(p_i q_j)^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}}.$$

Subtraindo 1 de ambos os membros da desigualdade acima e, em seguida, multiplicando pelo fator $A_\beta < 0$, temos

$$A_\beta\left[\left(\sum_{i,j}\pi_{ij}^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right] \leq A_\beta\left[\left(\sum_{i,j}(p_i q_j)^{\alpha}\right)^{\frac{\beta-1}{\alpha-1}} - 1\right],$$

e portanto, novamente a desigualdade de independência é satisfeita pela entropia H_n^*.

(iii) $\alpha > 1, \ \beta < 1$.

Assim, $A_\beta > 0$. Usando a convexidade da função $x \longmapsto x^{\alpha}$, $\alpha > 1$, temos que

$$\sum_j \frac{1}{m} q_{ij}^{\alpha} \geq \left(\sum_j \frac{1}{m} q_{ij}\right)^{\alpha}$$

para todo $i = 1, \ldots, n$, e portanto,

$$\sum_{i,j}(p_i q_{ij})^{\alpha} \geq \sum_{i,j}(p_i q_j)^{\alpha},$$

i.e,

$$\sum_{i,j} \pi_{ij}^{\alpha} \geq \sum_{i,j} (p_i q_j)^{\alpha}.$$

Elevando ambos os membros da desigualdade acima à potência $(\beta - 1)/(\alpha - 1) < 0$ e depois subtraindo 1, obtemos

$$\left(\sum_{i,j} \pi_{ij}^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}} - 1 \leq \left(\sum_{i,j} (p_i q_j)^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}} - 1.$$

Já que o fator A_β é positivo, multiplicando a desigualdade acima por este, obtemos novamente a desigualdade de independência para H_n^*, i.e., vale que

$$A_\beta \left[\left(\sum_{i,j} \pi_{ij}^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}} - 1 \right] \leq A_\beta \left[\left(\sum_{i,j} (p_i q_j)^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}} - 1 \right].$$

(iv) $\alpha > 1,\ \beta > 1.$

Aqui, temos que $A_\beta < 0$ e

$$\left(\sum_{i,j} \pi_{ij}^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}} - 1 \geq \left(\sum_{i,j} (p_i q_j)^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}} - 1.$$

Multiplicando a desigualdade acima pelo fator A_β, o qual é negativo, obtemos o resultado, isto é, a desigualdade de independência para H_n^* é, novamente, satisfeita para este caso.

Na próxima proposição provaremos a desigualdade de independência para canais com distribuições de probabilidades de entrada uniforme (i.e, as linhas da matriz de transição são permutações de um mesmo conjunto de m números).

Proposição 3.2.3

Seja $(q_{ij})_{n,m}$ a matriz cujas linhas são as permutações de um mes mo ... conjunto de m números. Então, a entropia de Mittal satisfaz a desigualdade de independência.

Demonstração: Seja $(a_1, a_2, \ldots, a_m)$ um tal conjunto de m números, e seja $\sum_{\ell=1}^{m} a_\ell^\alpha = A$. Então,

$$\sum_j q_{ij}^\alpha = \sum_{\ell=1}^{m} a_\ell^\alpha = A, \quad i = 1, \ldots, n;$$

$$\sum_j \sum_i p_i q_{ij}^\alpha = \sum_i \sum_j p_i q_{ij}^\alpha = \sum_i p_i . A = A_j$$

e portanto,

$$\sum_j q_{ij}^\alpha = \sum_j \sum_i p_i q_{ij}^\alpha \quad ; \quad i = 1, \ldots n.$$

Temos os seguintes casos a considerar:

(i) $0 \leq \alpha < 1$ e $0 \leq \beta < 1$.

Para os valores de α acima, temos

$$\sum_j \sum_i p_i q_{ij}^\alpha \leq \sum_j (\sum_i p_i q_{ij})^\alpha = \sum_j q_j^\alpha ,$$

isto é,

$$\sum_j q_{ij}^\alpha \leq \sum_j q_j^\alpha , \quad 0 \leq \alpha < 1.$$

Multiplicando a desigualdade acima por p_i^α e somando sobre i, obtemos

$$\sum_i \sum_j p_i^\alpha q_{ij}^\alpha \leq \sum_i \sum_j p_i^\alpha q_j^\alpha .$$

De $0 \leq \beta < 1$, temos que $A_\beta > 0$ e $(\beta - 1)/(\alpha - 1)$ é positivo. Dessa forma, temos que

$$A_\beta \left[\left(\sum_{i,j} \P_{ij}^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}} - 1 \right] \leq A_\beta \left[\left(\sum_{i,j} (p_i q_j)^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}} - 1 \right]$$

e a desigualdade de independência é satisfeita neste caso.

(ii) $0 < \alpha < 1$ e $\beta > 1$.

Aqui $A_\beta < 0$ e $(\beta - 1)/(\alpha - 1)$ é negativo. Novamente temos a desigualdade de independência satisfeita.

(iii) $\alpha > 1$ e $0 < \beta < 1$.

Neste caso, concluimos que

$$\sum_j q_{ij}^{\alpha} > \sum_j q_j^{\alpha}, \quad \alpha > 1,$$

que após multiplicarmos por p_i^{α} e somar sobre i, obtemos

$$\sum_{i,j} \P_{ij}^{\alpha} \geq \sum_{i,j} (p_i q_j)^{\alpha}$$

novamente a desigualdade de independência é satisfeita, uma vez que $A_\beta > 0$ e $(\beta - 1)/(\alpha - 1)$ é negativo.

(iv) $\alpha > 1$ e $\beta > 1$.

Obtemos que o fator A_β é negativo e, facilmente, vemos que $(\beta - 1)/(\alpha - 1)$ é positivo. Portanto, segue-se o resultado, i.e, a desigualdade de independência é novamente satisfeita para H_n^*.

Notemos que na demonstração da proposição acima, a propriedade de permutação-linha na matriz (q_{ij}) foi usada somente para termos a equação

$$\sum_j q_{ij}^{\alpha} = A \text{ (constante)},$$

para todo i = 1, ..., n.

Com efeito, a proposição supra é válida para uma distribuição de probabilidades condicional (q_{ij}), que satisfaz a equação acima, mas não necessariamente para a propriedade de permutação-linha.

Proposição 3.2.4

Se a distribuição de entrada (p_i) é uniforme, então a desigualdade de independência é satisfeita pela entropia de Mittal.

Demonstração: De fato, seja

$$L := \left[\sum_{i,j} (p_i \, q_{ij})^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}} = \left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} \left(\sum_{i,j} q_{ij}^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}}$$

e, seja

$$R := \left[\sum_{i,j} (p_i q_j)^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}} = \left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} \left(\sum_{i,j} q_j^{\alpha} \right)^{\frac{\beta-1}{\alpha-1}}$$

$$= \left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} \left[\sum_{j} (\sum_{i} q_j^{\alpha}) \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} \left[n \sum_{j} q_j^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} \left[n. \sum_{j} (\Sigma p_i \, q_{ij})^{\alpha} \right]^{\frac{\beta-1}{\alpha-1}}$$

$$= \left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} \left[n^{1-\alpha} \sum_{j} (\sum_{i} q_{ij}^{\alpha}) \right]^{\frac{\beta-1}{\alpha-1}}$$

(i) Para $0 \leq \alpha < 1$ e $0 \leq \beta < 1$, temos que

$$\sum_i \frac{1}{n} q_{ij}^{\alpha} \leq \left(\sum_i \frac{1}{n} q_{ij}\right)^{\alpha}, \quad \text{para todo } i = 1, \ldots, n;$$

(pela concavidade da função $x \longmapsto x^{\alpha}$, $0 \leq \alpha < 1$), que implica

$$\sum_{i,j} q_{ij}^{\alpha} \leq n^{1-\alpha} \sum_j \left(\sum_i q_{ij}\right)^{\alpha}$$

Elevando a desigualdade acima à potência $(\beta - 1)/(\alpha - 1)$, a qual é positiva, depois multiplicando pelo fator $\left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} > 0$, obtemos

$$\left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} \left[\sum_{i,j} q_{ij}^{\alpha}\right]^{\frac{\beta-1}{\alpha-1}} \leq \left(\frac{1}{n}\right)^{\alpha \frac{\beta-1}{\alpha-1}} \left[n^{1-\alpha} \sum_j \left(\sum_i q_{ij}^{\alpha}\right)\right]^{\frac{\beta-1}{\alpha-1}}$$

e portanto,

$$L \leq R.$$

Como o fator A_β é positivo, $0 \leq \beta < 1$, segue-se o resultado; quer dizer, a desigualdade de independência é satisfeita por H_n^*, neste caso.

(ii) $0 \leq \alpha < 1$ e $\beta > 1$.

Neste caso concluimos que $L \geq R$ e, facilmente, vemos que A_β é negativo. Portanto, novamente a desigualdade em questão se verifica para H_n^*.

Os outros casos podem ser provados de modo análogo.

Corolário 3.2.1

A desigualdade de independência é satisfeita pela entropia de Mittal se a distribuição de probabilidades conjunta (π_{ij}, $i=1,\ldots,n$, $j=1, \ldots, m$) é uniforme.

Demonstração: Uma distribuição de probabilidades conjunta uniforme ($\pi_{ij} = (1/mn)$, $\forall\, i,j$) resulta em uma distribuição de entrada (p_i) equiprovável. Com efeito,

$$p_i = \sum_j \pi_{ij} = \sum_{j=1}^{m} \frac{1}{nm} = m.\frac{1}{n.m} = \frac{1}{n}, \ \forall i.$$

Corolário 3.2.2

Se a distribuição de probabilidades condicional $(q_{ij},\ i=1, \ldots, n;\ j=1, \ldots, m)$ é uniforme, então a entropia de Mittal satisfaz a desigualdade de independência.

Demonstração: Já que a distribuição de probabilidade condicional é uniforme, temos que a propriedade de permutação-linha é satisfeita, e portanto, resulta numa distribuição de probabilidades de saída (q_j) uniforme, i.e, temos

$$q_j = \sum_i \pi_{ij} = \sum_i p_i\, q_{ij} = \sum_i p_i \frac{1}{m} = \frac{1}{m} \cdot \sum_i p_i = \frac{1}{m}$$

Pela proposição 3.2.2, segue-se o resultado.

3.2.1. Caso Particular

Como vimos (seção 1.4.5), fazendo $\alpha = 1/\gamma$ e supondo que $\beta = 2-\gamma$, na entropia de Mittal, H_n^*; esta reduz-se ao caso particular.

$$_\gamma H_n\left(p; \frac{1}{\gamma}, 2-\gamma\right) = \frac{1}{2^{\gamma-1}-1}\left(\left(\sum_{i=1}^{n} p_i^{1/\gamma}\right)^{\gamma} - 1\right), \quad \gamma > 0, \quad \gamma \neq 1,$$

a qual chamamos de entropia gama generalizada. Como tal, esta medida de informação também satisfaz a desigualdade de independência, considerando-se as classes de distribuições de probabilidades aqui estudadas.

REFERÊNCIAS BIBLIOGRÁFICAS

|1| ACZÉL, J., and DARÓCZY, Z: "On measures of Information and their Characterizations", Academic Press, New York (1975).

|2| ASH, R: Information Theorg, John Wiley and Sons, New York (1965).

|3| BRILLOWIN, L: Science and Information Theory, Academic Press, New Yokr, (1956).

|4| BOEKEE, D.E, and VAN DER LUBBE, J.C.A: The R-Norm Information Measure, Inform. Contr., 45, 136-155, (1980).

|5| DARÓCZY, Z: On Shannon Measures of Information, Magyar, Trad. Acad. Mat. Fiz. Qrzt, Kőzl, 19, 9-24, (1969).

|6| DARÓCZY, Z: Generalized Information Function, Inform. Contr., 16, 36-51.

|7| EL-SAYED, ABU-BAKR: The Independence Inequality and its Applications to Information Theory, Inform. Contr., 35, 229-245, (1977).

|8| FEINSTEIN, A: "Foundations of Information Theory", McGraw-Hill, New York/London/Toronto (1958).

|9| FORTE, B. and DARÓCZY, 2: A Characterization of Shannon's Entropia, Bell. U.M.I., 4, 631-635, (1968).

|10| FORTE, B. and Ng, C.T: On a Characterization of Entropias of Degrees α, Utilitas Mathematics, 4, 193-205, (1973).

|11| HARDY, G.H, LITTLEWOOD, J.E. and POLYA, G: Inequalities, Cambridge University Press, (1967).

|12| HARTLEY, R.V.L: Transmission of Information, Bell System Tech. J., 7: 535 (Julho, 1928).

|13| JELINEK, F: Probabilistic Information Theorg, McGraw-Hill, Genes in System Science, (1968).

|14| KULLBACK, S: Information Theory and Statistics, John Wiley and Sons, New York (1967).

|15| NYQUIST, H: Certain Factors Affecting Telegraph Speed, Bell System Tech. J. 3-324 (abril, 1924).

|16| NYQUIST, H: Certain Topics in Telegraph Transmission Theory, AIEE Trans., 47-617 (abril, 1928).

|17| PINTACUDA, N: Shannon Eutropia. A Move General Derivation, Statistics, 26, 509-524, (1966).

|18|, QUASTLER, H: Information Theory and Biology, University of Illinois Press, Urbana, (1953).

|19|, RÉNYI, A: On Measures of Entropy and Information, Proc. Fourth Berkeley - Sym p. Math. Statistic and Probability, University of California Press, 1, 547.561, (1970).

|20|, RÉNYI, A: Probability Theory, North-Holland, Amsterdam, (1970).

|21|, SHARMA, B.D. and TANEJA, I.J: Three-Generalized-Additive Measures of Entropia, EIK 13, 419-433 (1977).

|22|, SHARMA, B.D., and TANEJA, I.J: Entropy of Type (α,β) and Their Generalized Measures in Information Theory, Metrika, 22, 205-215, (1975).

|23|, SHARMA, B.D., and MITTAL, D.P: New Non-additive Measures of Entropy for Disciete Probability Distribuitions, Journal of Math. Science, 10, 28-40, (1975).

|24|, TANEJA, I.J: A Joint Characterization of Shannon's Entropy and Entropy of Type α throwgh Functional Eq. - J. Math. Sci., 10, 69-74, (1975).

|25|, TANEJA, I.J: Some Contribuitions to Information Theory-I (A Survey). On Measures of Information, J. Comb. Inform. and System. Sci., 4, 259-280, (1979).

|26|, THEIL, H: Economics and Information Theory, North-Holland, Amsterdam(1967).

|27|, VAN DER PYL, T: Information d'ordre α et de type β : Axiomatique, Proprietes, Thése de 3º cycle, Paris (1976).

|28|, WIENER, N: "Cybernetics", The Technology Press of the Massachusetts Institute of technology and John Wiley and Sons, Inc., New York, (1948).

|29|, ZATELLI, A: Entropia Paramétrica Generalizada e a Probabilidade de Erro, Tese de Mestrado, UFSC (1982).